

# Determinants of Substrate Specificity in the NS3 Serine Proteinase of the Hepatitis C Virus

Jan Oliver Koch and Ralf Bartenschlager<sup>1</sup>

*Institute for Virology, Johannes-Gutenberg University, Obere Zahlbacher Strasse 67, 55101 Mainz, Germany*

*Received May 27, 1997; returned to author for revision July 14, 1997; accepted August 1, 1997*

Processing of the nonstructural polyprotein of the hepatitis C virus (HCV) requires the serine-type proteinase located in the amino-terminal domain of NS3. To identify residues within NS3 determining substrate specificity, a mutation analysis was performed. Using sequence alignments and three-dimensional structure predictions, amino acids assumed to be important for specificity were replaced and the enzymes were tested in an intracellular *trans*-processing assay for their effects on cleavage of an NS4B-5B substrate. For some of the substitutions at positions 133, 134, 135, 136, 138, 152, 155, 157, and 169, slightly reduced processing efficiencies were observed but in no case was the substrate specificity altered. In contrast, substitutions of the phenylalanine at position 154 resulted in a modified cleavage pattern, suggesting an important role for this residue in substrate specificity. To substantiate this assumption, a panel of NS4B-5B substrates carrying different P1 residues at the NS4B/5A site were tested for cleavage by these altered proteinases. We found that substitution of Phe-154 by alanine, by valine, and particularly by threonine generated enzymes with the following affinities for aliphatic P1 residues: C > L > I > V for 154 F → A, C = L > I > V for 154 F → V and L > C > I > V for 154 F → T. Neither leucine nor isoleucine nor valine was accepted by the parental NS3 proteinase, showing that Phe-154 is an important determinant for substrate specificity. Furthermore, we present evidence that Ala-157 plays an additional but minor role for this property.

© 1997 Academic Press

## INTRODUCTION

The hepatitis C virus (HCV) is one of the major causes of transfusion-associated hepatitis worldwide. Most, if not all, infections become chronic and result in various clinic outcomes, including acute hepatitis, chronic hepatitis, and liver cirrhosis. Chronic HCV infections, in particular those with cirrhosis, are also associated with the development of hepatocellular carcinoma (for reviews see Cuthbert, 1994; Houghton, 1996). No effective therapy nor any vaccine is yet available.

The hepatitis C virus is classified together with the flaviviruses and the animal-pathogenic pestiviruses in the family *Flaviviridae* (Murphy *et al.*, 1995). These viruses have in common a single-stranded RNA genome of positive polarity, which in the case of HCV has a length of about 9600 nucleotides. It carries a single long open reading frame, encoding a 3010- to 3033-amino-acid-long polyprotein (for reviews see Rice, 1996; Bartenschlager, 1997). It is cleaved co- and posttranslationally into at least 10 mature viral proteins, which are arranged within the polyprotein as follows (from the amino to the carboxy terminus): H<sub>2</sub>N-C-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B-COOH (for reviews see Rice, 1996; Bartenschlager, 1997). For several of these proteins, distinct functions can be ascribed. The core protein (C), a basic RNA-binding molecule, most likely is the major constituent of the

capsid (Hsu *et al.*, 1993; Santolini *et al.*, 1994; Matsumoto *et al.*, 1996). E1 and E2 are the viral envelope glycoproteins (Matsuura *et al.*, 1992; Spaete *et al.*, 1992; Ralston *et al.*, 1993; Dubuisson *et al.*, 1994). Nonstructural protein (NS) 2 most likely encodes the enzymatically active domain of the NS2-3 proteinase (Grakoui *et al.*, 1993a; Hijikata *et al.*, 1993). NS3 is a bifunctional molecule with the amino-terminal domain carrying a serine-type proteinase (Bartenschlager *et al.*, 1993; Eckart *et al.*, 1993; Grakoui *et al.*, 1993b; Hijikata *et al.*, 1993; Tomei *et al.*, 1993) and the carboxy-terminal domain a nucleoside triphosphatase/RNA helicase activity (Suzich *et al.*, 1993; Kim *et al.*, 1995). NS4A is an NS3 proteinase cofactor (Bartenschlager *et al.*, 1994; Failla *et al.*, 1994; Lin *et al.*, 1994; Tanji *et al.*, 1995) and NS5B was shown to be an RNA-dependent RNA polymerase (Behrens *et al.*, 1996; Lohmann *et al.*, 1997). No functions have yet been ascribed to NS4B and NS5A.

Proteolytic processing of the viral polyprotein is mediated by two classes of proteinases: the host cell signal peptidases cleaving the C-NS2 region and two viral enzymes responsible for processing of the remainder. These are the NS2-3 proteinase, a zinc-dependent metalloproteinase cleaving at the NS2-3 junction (Grakoui *et al.*, 1993a; Hijikata *et al.*, 1993), and the NS3 proteinase located in the amino-terminal 181 residues of the NS3 protein, which is responsible for processing at all other sites in the NS region (Bartenschlager *et al.*, 1993; Eckart *et al.*, 1993; Grakoui *et al.*, 1993b; Hijikata *et al.*, 1993; Tomei *et al.*, 1993). The scissile bond at the NS3/4A junction is cleaved most likely by an intramolecular (*cis*)

<sup>1</sup> To whom correspondence and reprint requests should be addressed. Fax: 0049-6131-395604. E-mail: bartnsch@mzdmza.zdv.uni-mainz.de.

reaction, whereas processing at the NS4A/4B, NS4B/5A, and NS5A/5B sites can be mediated *in trans* (Tomei *et al.*, 1993; Bartenschlager *et al.*, 1994; Failla *et al.*, 1994; Lin *et al.*, 1994). Cleavage activity of the NS3 proteinase is greatly enhanced by the NS4A cofactor. Both proteins form a stable complex, for which the amino-terminal 22 residues of NS3 and the central NS4A region are required (Bartenschlager *et al.*, 1995b; Failla *et al.*, 1995; Lin *et al.*, 1995; Satoh *et al.*, 1995; Kim *et al.*, 1996). Although this association is essential for cleavage at the NS3/4A, NS4A/4B, and NS4B/5A sites, processing between NS5A and NS5B can be mediated by the NS3 proteinase in the absence of NS4A, albeit with a low efficiency (Bartenschlager *et al.*, 1994; Failla *et al.*, 1994; Lin *et al.*, 1994).

The substrate requirements of the NS3 proteinase have been studied intensively. A consensus sequence could be deduced from sequence alignments of the four cleavage sites within the polyprotein (Grakoui *et al.*, 1993b; Pizzi *et al.*, 1994). It reads Asp/Glu-X-X-X-Cys/Thr ↓ Ser/Ala, where X represents any amino acid. While a cysteine residue is present at all P1 positions of *trans*-cleavage sites, a threonine residue is found at the NS3/4A *cis*-cleavage site (amino acids extending from the cleaved bond toward the amino terminus are denoted P1, P2, P3, etc., and those extending toward the carboxy terminus are designated P1', P2', P3', etc.; Schechter and Berger, 1967). Using site-directed mutagenesis experiments, it was shown that the P1 residue is most sensitive to mutations and is therefore a major determinant of specificity within the substrate whereas the conserved P6 and P1' residues play only a minor role (Kolykhalov *et al.*, 1994; Komoda *et al.*, 1994; Leinbach *et al.*, 1994; Bartenschlager *et al.*, 1995a).

Molecular modeling studies have been used to predict determinants of substrate specificity within the NS3 proteinase domain (Pizzi *et al.*, 1994). They suggested that the phenylalanine residue at position 154 of NS3 is a major determinant shaping the substrate-binding pocket (SBP) accommodating the P1 residue (the S1-binding pocket). This assumption was substantiated with a limited mutation analysis showing that two distinct substitutions within the putative SBP led to an alteration of NS3 substrate specificity (Failla *et al.*, 1996). Very recently, the three-dimensional structures of the NS3 proteinase domain (Love *et al.*, 1996) and the proteinase domain complexed with a synthetic NS4A peptide were determined (Kim *et al.*, 1996). These analyses show that the SBP is shallow, nonpolar, and determined primarily by residues Leu-135, Phe-154, and Ala-157. It was suggested that Phe-154 forms the bottom of the SBP and limits the length of the P1 side chain fitting into the S1 pocket. In contrast, the contributions of Leu-135 and Ala-157 for specificity are less clear.

To further study the role these three residues play for NS3 substrate specificity and to determine whether residues in close proximity might contribute to specificity,

we have used an intensive mutation analysis. Several amino acid substitutions were introduced at these positions and the substrate specificities of the proteinases were determined in an intracellular *trans*-cleavage assay using NS4B-5B substrates carrying various P1 residues at the NS4B/5A site. We found that among the single substitutions, only those affecting Phe-154 substantially broadened specificity whereas those of Leu-135 and Ala-157 had no effect. However, substitutions of Ala-157 contributed to an alteration of specificity when combined with exchanges of Phe-154.

## MATERIALS AND METHODS

### Cells and viruses

BHK-21 and human TK<sup>-</sup> 143 cell monolayers were grown in Dulbecco's modified minimal essential medium (DMEM) supplemented with 2 mM L-glutamine, nonessential amino acids, 100 U penicillin per milliliter, 100 µg streptomycin per milliliter, and 10% fetal calf serum (FCS). vTF7-3, a recombinant vaccinia virus expressing the RNA polymerase of the T7 bacteriophage was described by Fuerst *et al.* (1986). Virus stocks were grown in TK<sup>-</sup> 143 cell monolayers. Virus titers were measured by plaque assay on the same cell line.

### Antisera

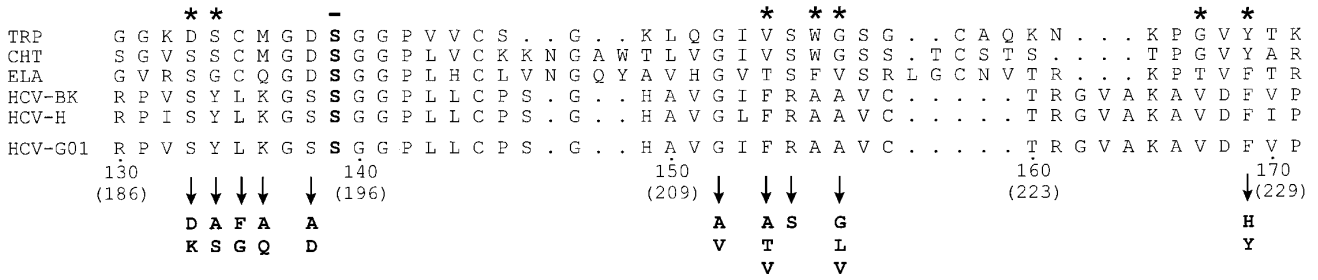
Monospecific antisera directed against NS3 (amino acids 1007–1246), NS5A (amino acids 2102–2235), and NS5B (amino acids 2421–2626) have been described recently (Bartenschlager *et al.*, 1994, 1995b). The monoclonal antibody directed against the hemagglutinin (HA) epitope (CYPYDVPDYASL) was purchased from Boehringer Mannheim (Germany).

### Plasmid constructions

The basic vectors pTM1-2 and pHA and the constructs pTM1027-1238, pTM1659-1711, and pTM1712-3010 have been described previously (Bartenschlager *et al.*, 1995b; Koch *et al.*, 1996; Fig. 1).

Standard recombinant DNA techniques were used for construction of all plasmids (Sambrook *et al.*, 1989). Site-directed mutagenesis was done by PCR methodology as described recently (Bartenschlager *et al.*, 1995b; Koch *et al.*, 1996). In case of mutations of the proteinase, the 642-nucleotide-long PCR fragments were inserted via *Nco*I and *Xba*I into pTM1-2 restricted with *Nco*I and *Spe*I. In case of the mutations at the NS4B/5A site, the 718-nucleotide-long PCR fragments were inserted via *Hpa*I and *Eco*RI sites into pTM1712-3010. All subcloned PCR fragments were sequenced using IRD-41-labeled primers and a LiCoR Model 4000 DNA sequencer (Li-CoR, Lincoln, NE). Subsequently the complete NS4B-5B fragments were recloned into pHA via the *Nco*I and *Spe*I restriction sites. The resulting constructs allowed the ex-

**A**



**B**

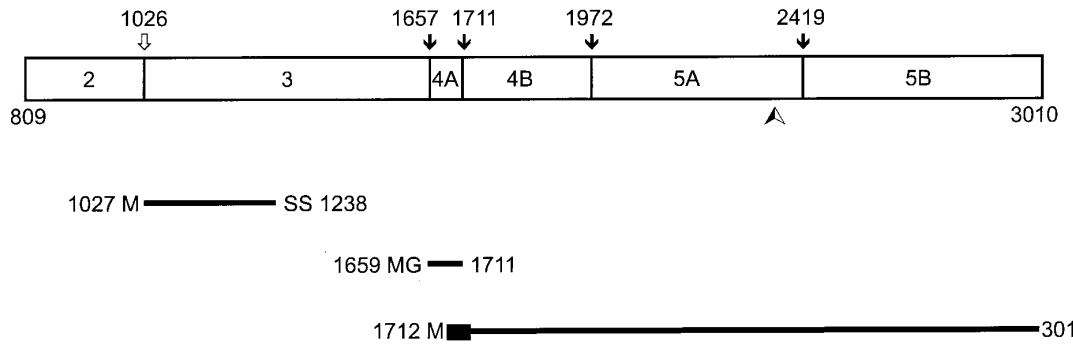


FIG. 1. (A) Alignment of the NS3 region spanning the amino acids around the putative SBP. Sequences corresponding to our HCV isolate (HCV-G01), HCV isolates of genotypes 1a (HCV-H; Inchauspe *et al.*, 1991), and 1b (HCV-BK; Takamizawa *et al.*, 1991) and the cellular serine-type proteinases trypsin (TRP), chymotrypsin (CHT), and elastase (ELA) were aligned according to the report of Bazan and Fletterick (1988). The active site serine residue is marked with a line; amino acids important in shaping the SBP of the cellular proteinases are labeled with an asterisk. Gaps were inserted to match highly conserved residues. Amino acids selected for site-directed mutagenesis are given underneath with arrows pointing to the substituting residues. Numbers below the alignment refer to amino acids of NS3 of the G01 isolate (same numbering as HCV-BK) or to the chymotrypsin numbering (given in parentheses). (B) Detailed view of the NS region of HCV (amino acids 809–3010). Arrows indicate the cleavage sites for the NS2-3 (♠) and the NS3 (♣) proteinases with numbers above the arrows referring to the corresponding P1 residues. The arrow close to the end of NS5A (▼) indicates the approximate position of the alternative cleavage site. Sequences expressed with the individual constructs are drawn below as lines. Numbers beside the lines refer to the first and last amino acids of the expressed sequence and the letters refer to heterologous amino acids at the amino and carboxy termini of the expressed proteins. The HA epitope is indicated by a box. The nomenclature for each polypeptide is given to the left.

pression of NS4B-5B polypeptides carrying the HA epitope at their amino termini (Fig. 1).

### *Trans*-cleavage assay using transient expression with the vaccinia virus T7 hybrid system and detection of cleavage products

A total of  $7.5 \times 10^4$  BHK 21 cells were seeded in a 15-mm-diameter dish 20 hr prior to transfection. Cells were infected with the vTF7-3 vaccinia virus at a multiplicity of infection of 5 for 1 hr at room temperature. The inoculum was replaced by DMEM containing 10% FCS and cells were incubated at 37° for 30 min. After washing twice with Optimem (Gibco BRL, Eggenstein, Germany) a total of 0.35  $\mu$ g DNA corresponding to a mixture of equal amounts of three plasmids was introduced into the cells using lipofectamine (Gibco BRL) according to the

instructions of the manufacturer. After 3 hr of incubation, cells were washed twice with methionine-free DMEM without FCS, overlaid with DMEM containing 100  $\mu$ Ci/ml Express protein labeling mixture (NEN Life Science, Cologne, Germany), and incubated for 3 hr at 37°. Cells were washed once with phosphate-buffered saline and lysed in NPB (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Nonidet-P40, 1% sodium desoxycholate, 0.1% SDS) supplemented with 1 mM phenylmethylsulfonyl fluoride and 0.001 trypsin inhibitor units per milliliter aprotinin (Sigma, Deisenhofen, Germany). For isolation of HCV-specific proteins by immunoprecipitation (IP), cell lysates were cleared by 10 min of centrifugation at 15,000 *g* and one-quarter of the supernatant was used for isolation of NS3/4A complexes by IP under nondenaturing conditions (Koch *et al.*, 1996). To reduce the background due to nonspecific binding of cellular or vaccinia virus proteins,

IPs were performed with the remainder of the lysate under denaturing conditions as described recently (Bartenschlager *et al.*, 1994). Immunocomplexes were analyzed by SDS–11% polyacrylamide gel electrophoresis (PAGE) or, in the case of the NS3/4A complex by SDS–Tricine–11% PAGE. Gels were stained with Coomassie blue, treated for fluorography with sodium salicylate, and exposed to Kodak BioMax MR film (Sigma). In some cases stained gels were directly exposed to BioMax MS films with an intensifying screen (TranScreen, Integra Biosciences, Fernwald, Germany). Quantifications of HA-4B proteins were done by densitometry scanning and the amounts of coexpressed NS3 and NS5B proteins were used as internal controls to correct for transfection efficiencies.

## RESULTS

### Experimental strategy

To identify residues which might contribute to substrate specificity, the amino acid sequence of the HCV NS3 proteinase domain of our isolate (G01) and of two other isolates corresponding to genotypes 1a and 1b (HCV-H and HCV-BK, respectively) were aligned with the sequences of well-characterized cellular serine proteinases (trypsin, chymotrypsin, and elastase; Fig. 1A). Furthermore, we included the model described by Pizzi and co-workers (1994) suggesting that Phe-154 is a major determinant for substrate specificity. Finally, while this work was in progress, two independent groups reported the determination of the three-dimensional X-ray crystal structure of the NS3 proteinase (Kim *et al.*, 1996; Love *et al.*, 1996). They showed that the shape of the SBP is determined primarily by Leu-135, Phe-154, and Ala-157 (chymotrypsin numbering 191, 213, and 216, respectively; Fig. 1A). Based on this information, we selected several residues clustered around positions 135 and 154 as targets for site-directed mutagenesis. The substitutions were introduced into a proteinase corresponding to the first 211 amino acids of NS3 (Fig. 1B). This enzyme, when coexpressed with the NS4A cofactor, was shown to be fully competent for cleavage at all NS3-dependent sites (Bartenschlager *et al.*, 1994). Proteinase and cofactor were expressed with the vaccinia virus/T7 hybrid system in BHK-21 cells together with an NS4B-5B substrate. This substrate was selected for several reasons: (i) It contained two cleavage sites, an NS4A-dependent site (NS4B/5A) and a site at which the NS3 proteinase could cleave in the absence of cofactor (NS5A/5B) albeit with a low efficiency. Therefore, this substrate allowed us to determine whether any of the substitutions had a differential effect on cleavage at these two sites. (ii) Based on a recent report, we anticipated that the substitutions would not change specificity completely but rather broaden it (Failla *et al.*, 1996). Since the substrate contained two cleavage sites it allowed us to determine the effect of P1 substitutions at one site whereas cleavage

at the other, unaltered site could be used as an internal control for overall enzymatic activity. (iii) NS4B-5B is a long protein, increasing the probability of containing alternative cleavage sites which might be recognized by altered NS3 proteinases. Thus, an altered cleavage pattern could be taken as strong evidence for a modification of NS3 substrate specificity. It should be noted that due to different expression levels and/or stabilities of expressed proteins and processing products, this experimental system gave only qualitative information. Therefore, we did not attempt to quantify enzymatic activities of the different proteins in terms of  $K_m$  or  $k_{cat}$ .

To facilitate detection of NS4B, an HA epitope was fused to the amino terminus (Fig. 1B), allowing immunoprecipitation of NS4B with an HA-specific antiserum. To reduce the background of cellular or vaccinia virus proteins binding nonspecifically to protein A–Sepharose or immunoglobulin, immunoprecipitations with antisera directed against the HA epitope, NS5A or NS5B, were done under denaturing conditions. Immunoprecipitations with the NS3-specific antiserum were done under conditions allowing the coprecipitation of NS3 proteinase with NS4A and therefore allowing determination of whether the introduced substitution affected the interaction with the cofactor. The amount of NS3 detected was used to correct for different expression levels of proteinases obtained with the individual plasmids.

### Essential role of Phe-154 in substrate specificity

The result of an experiment in which all proteinases were tested for coprecipitation with the NS4A cofactor and cleavage of the HA4B-5B substrate is shown in Fig. 2. For the unaltered proteinase (wt, lane 29), the precursor was almost completely cleaved into HA4B, NS5A, and NS5B and only very small amounts of unprocessed HA4B-5B and HA4B-5A intermediates were detected. Cleavage was specific because no products were found with an NS3 proteinase in which the active site serine residue was replaced by alanine (3A, lane 30). For unknown reasons, the NS3 proteinase domain appeared as a double band. This observation did not correlate with any of the experimental parameters and was not reproducible (see Fig. 4).

As deduced from the increased amounts of HA4B-5B precursor, which can best be seen in the immunoprecipitation with the 5A-specific antiserum, all substitutions affected cleavage efficiencies to various extents. The most drastic effect was found for the 135 L → G and the 152 G → V substitutions which completely inactivated the proteinases (lanes 6 and 12). The low expression levels of enzymes could not account for the lack of precursor cleavage because a similarly low expression level was obtained, e.g., with the 3<sub>169H</sub> proteinase which was clearly active (lane 20). A less detrimental effect was found for the 157 A → L and the 157 A → V substitutions (lane 18, 19) whereas the glycine substitution at the same position

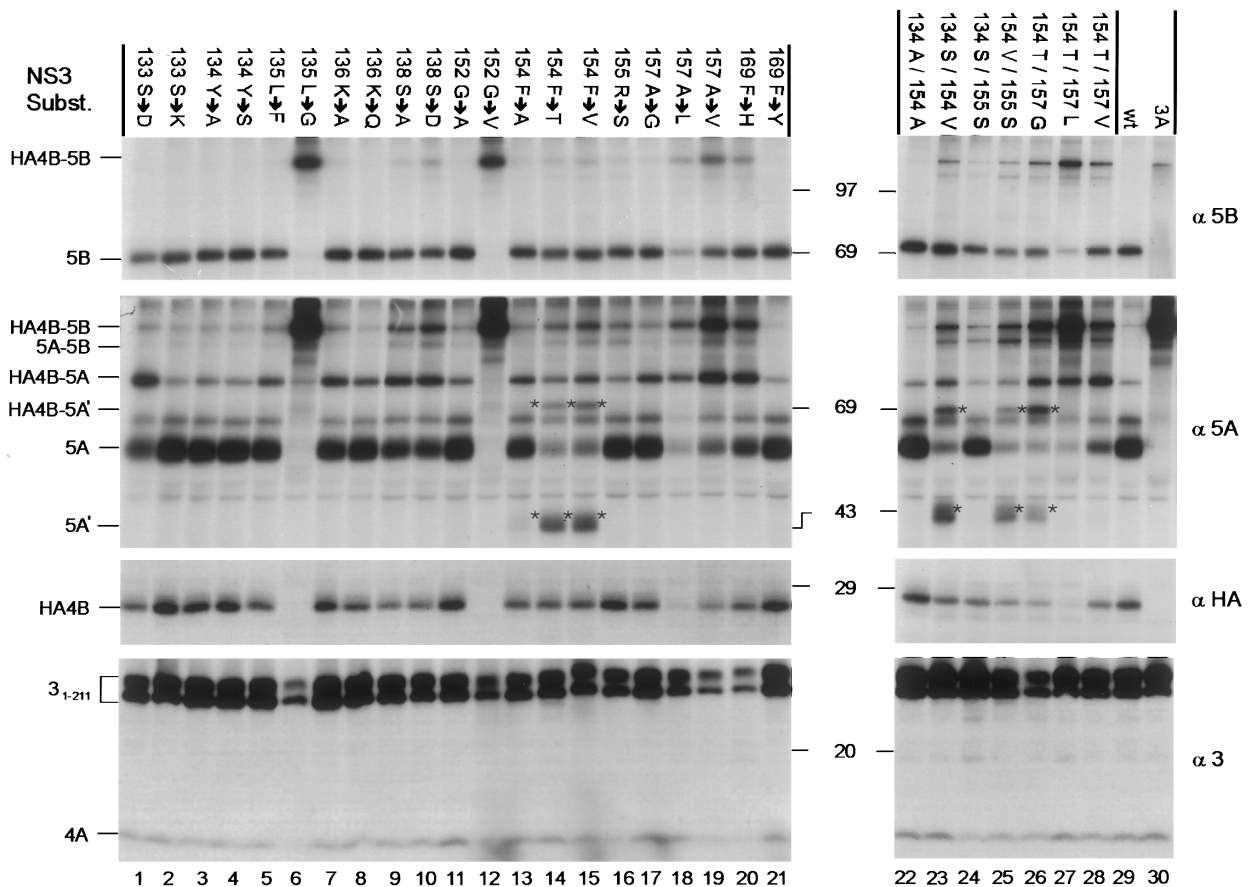
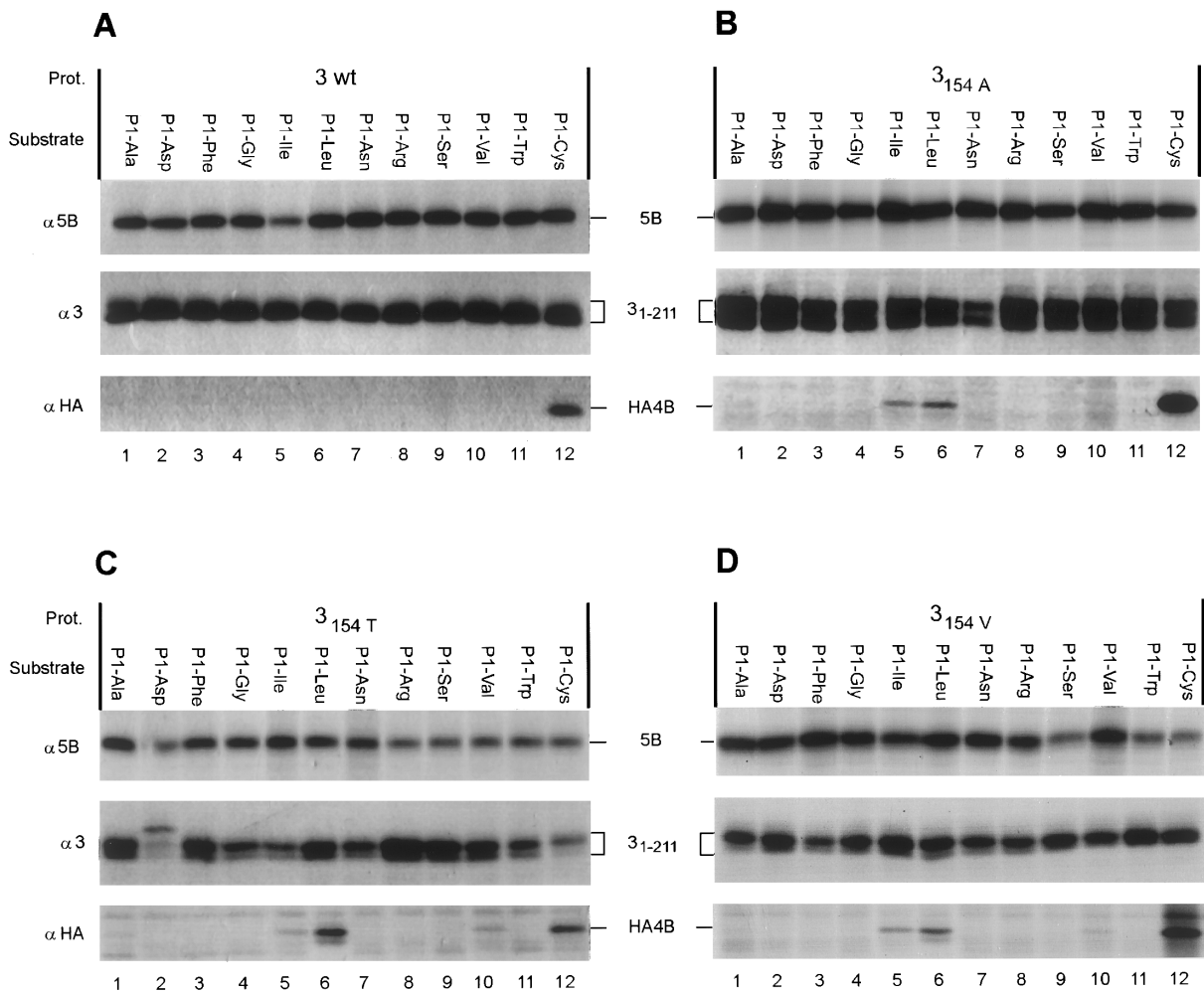


FIG. 2. Substitutions within NS3 and their effects on cleavage activities and specificities. Cells infected with vTF7-3 were transfected with plasmids directing the expression of NS4A<sub>2-54</sub>, the HA4B-5B substrate, and a proteinase specified above each lane. After 3 hr of labeling cells were lysed and HCV-specific proteins were isolated by immunoprecipitation using antisera ( $\alpha$ ) directed against NS5B, NS5A, or the HA epitope under denaturing conditions or, in case of the NS3/4A complex, with an NS3-specific antiserum under native conditions. Proteins were analyzed by SDS-11% PAGE or by SDS-Tricine-11% PAGE (for the NS3/4A complex) and detected by fluorography. Results obtained with single substitutions are shown on the left, and those obtained with double substitutions are shown on the right. HCV-specific proteins are identified to the left of the gels. The products generated by cleavage at the alternative site within NS5A are marked with asterisks. Numbers between the panels refer to the molecular masses of protein size standards (in kilodaltons). Antisera used for immunoprecipitations are specified to the right. The results obtained with the parental proteinase (wt) and the mutant in which the active-site serine residue was replaced by an alanine (3A) are shown in lanes 29 and 30, respectively.

had no effect (lane 17), indicating that a small hydrophobic residue at this position is required for full proteolytic activity. In the case of substitutions of Phe-169 a significant reduction of cleavage activity was found when a histidine residue was introduced (lane 20) whereas substitution by tyrosine had no effect (lane 21), indicating the requirement for a large hydrophobic (aromatic) amino acid at this position.

The only substitutions for which an altered cleavage pattern was obtained were those affecting the phenylalanine residue at position 154. In all cases, in addition to the natural cleavage products, two proteins with apparent molecular weights of approximately 70 and 43 kDa were found (lanes 13-15). While the smaller protein reacted only with the 5A-specific antiserum, the 70-kDa protein could be precipitated with both the 5A- and the HA-specific antisera (not shown). These immune reactivities and the apparent molecular masses suggested that the proteins were generated most likely by cleavage at

an additional site close to the carboxy terminus of NS5A, removing an approximately 15-kDa fragment from NS5A or the HA4B-5A precursor (Fig. 1B). Consequently these proteins were designated 5A' and HA4B-5A'. Since these cleavage products were not produced by the parental proteinase, this result suggested that the proteinases carrying substitutions of Phe-154 had an altered substrate specificity. Furthermore, the size of the substituting residue was important. While the 3<sub>154A</sub> proteinase behaved almost like the parental enzyme (lane 13), introduction of bulkier residues at position 154 increased cleavage efficiency at the alternative site (3<sub>154T</sub> and 3<sub>154V</sub>; lanes 14 and 15). It should be noted that the same results were obtained with an NS4B-5B substrate, demonstrating that the foreign HA sequence at the amino terminus of NS4B had no effect on cleavage efficiency and substrate specificity (not shown). In summary, these data demonstrate that Phe-154 of the NS3 proteinase is a major determinant of substrate specificity.



**FIG. 3.** Substrate specificity of proteinases carrying various substitutions of Phe-154. vTF7-3-infected cells were transfected with plasmids allowing the expression of NS4A<sub>2-54</sub>, together with the parental NS3 proteinase (A), or one of the given proteinase mutants (B–D) and HA4B-5B substrates carrying various substitutions of the P1 residue at the NS4B/5A cleavage site. As a control, all proteinases were also tested for cleavage of the unaltered substrate (P1-Cys; lane 12 in each panel). At the end of the 3-hr labeling period cells were lysed and after immunoprecipitation with antisera directed against NS3, NS5B, or the HA epitope, precipitated proteins were analyzed by SDS–PAGE (NS5B and HA4B) or SDS–Tricine–PAGE (NS3).

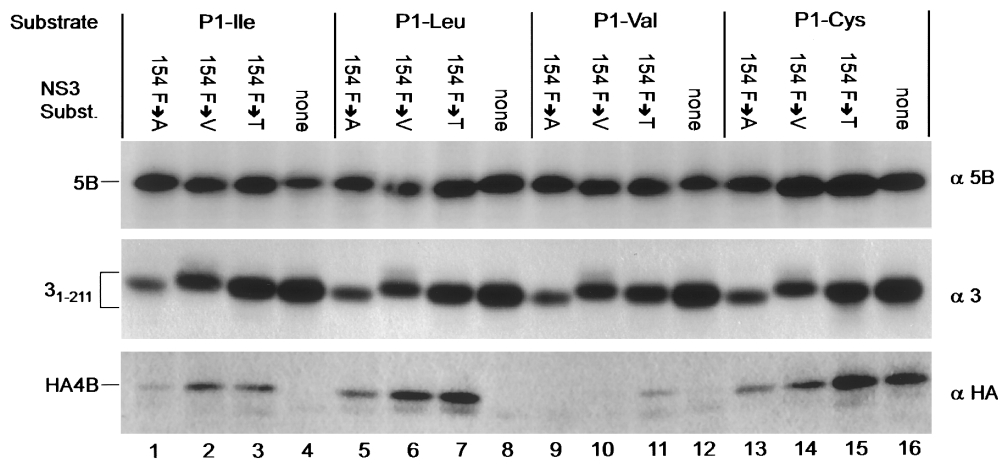
To analyze whether the simultaneous substitution of two residues assumed to be important for substrate specificity might have a synergistic effect, a panel of double mutants was generated and tested in the same way (Fig. 2, right). Focusing on Phe-154, substitutions at this position were combined with several other exchanges of amino acids which might contribute to specificity. As exemplified for the 134 Y → S/154 F → V double mutant, in most cases two substitutions with one affecting Phe-154 behaved like the single substitution at this position (compare lane 23 with lane 15). The exception was the double mutation 154 F → T/157 A → L, which drastically reduced activity (lane 27). The reason for this effect will be described below.

#### Cleavage activities and substrate specificities of proteinases substituted at position 154

To characterize the substrate specificities of the substitutions at position 154 more precisely, we generated

a panel of HA4B-5B substrates in which the P1 cysteine residue of the NS4B/5A cleavage site was altered (Fig. 3). We anticipated that an NS3 proteinase with an altered S1 pocket should accommodate P1 residues which are not or only very poorly accepted by the parental proteinase. Since the processing pattern of NS5A was too complex to allow quantifications, we used the amount of HA4B produced during a 3-hr labeling period to determine efficiency of cleavage at the altered site.

In the first set of experiments we analyzed the substrate specificity of the parental NS3<sub>1-211</sub> proteinase using 12 substrates, each with a different P1 residue at the NS4B/5A junction. As shown in Fig. 3A, only the P1-Cys (wild type) substrate was cleaved efficiently (lane 12). With the exception of the P1-Ala substitution where a very low cleavage was detected (barely visible on the film), all other P1-residues were not accepted. This result, which agrees with two recent reports (Komoda *et al.*, 1994; Bartenschlager *et al.*, 1995a), is at variance with



**FIG. 4.** Comparison of cleavage efficiencies of the parental NS3 proteinase with proteinases carrying threonine, valine, or alanine substitutions at position 154. Cells were infected with vTF7-3 and transfected with plasmids directing the expression of NS4A<sub>2-54</sub>, one of the proteinases given above the lanes and HA4B-5B substrates in which the P1 residue of the NS4B/5A site was unchanged (lanes 13–16) or replaced by isoleucine (lanes 1–4), leucine (lanes 5–8), or valine (lanes 9–12). HCV-specific proteins are specified on the left side of the gels; antisera used for immunoprecipitations are specified on the right. All proteins were analyzed by SDS–11% PAGE.

the results described by Kolykhalov and co-workers (1994) but can be explained by different levels of sensitivity, different stabilities of the NS4B proteins, or different HCV sequences. When the proteinases carrying substitutions of Phe-154 were tested in the same way, a different pattern of specificity was found. In all cases, in addition to the P1-Cys residue, isoleucine, leucine and, to a lesser extent, valine were accepted as P1 amino acids, demonstrating that the proteinases had an altered substrate specificity and accepted amino acids with large nonpolar side chains (Figs. 3B–3D, lanes 5, 6, 10, and 12; the nature of the additional band in the NS3-specific immunoprecipitation shown in Fig. 3C, lane 2, is not known but in repetitive experiments it was not reproducible).

To quantify the relative alterations of substrate specificities of these enzymes, several transfections were performed in which proteinases with an altered cleavage pattern were compared side by side for processing of HA4B-5B substrates carrying P1-Leu, -Ile, -Val, or -Cys residues at the NS4B/5A junction. As shown in Fig. 4 the efficiencies with which these P1 residues were accepted depended on the substituting residue in the S1 pocket. Using densitometry scanning to determine the amount of HA4B [corrected for the expression levels of NS3 proteinase and substrate (NS5B)], we found that the 3<sub>154A</sub> proteinase cleaved at the P1-Cys most efficiently (Table 1). In the case of the 3<sub>154V</sub> proteinase, leucine and cysteine were accepted equally well whereas the 3<sub>154T</sub> enzyme cleaved at the P1-Leu with the highest efficiency (Table 1). In summary, for the different proteinases the following orders of affinities for P1 residues were found: Cys > Leu > Ile > Val for the 154 F → A exchange, Cys = Leu > Ile > Val for the 154 F → V substitution, and Leu > Cys > Ile > Val for the F → T exchange (Table 1). The fact that P1-Ala was not accepted by these proteinases and P1-Val was accepted only very poorly,

whereas P1-Leu and P1-Ile were rather good P1 residues, suggested that we had generated enzymes with opened SBPs which interacted most favorably with P1 residues containing large aliphatic side chains.

**Minor roles of residues Leu-135 and Ala-157 in substrate specificity**

Based on the three-dimensional structure of the NS3 proteinase, it was suggested that the shape of the SBP is determined primarily by Leu-135, Phe-154, and Ala-157 (Kim *et al.*, 1996; Love *et al.*, 1996). In the experiment shown in Fig. 2 we found that substitution of Leu-135 by phenylalanine had no effect on specificity whereas substitution by glycine completely blocked cleavage at the P1-Cys residue. To analyze these enzymes in more detail they were tested for cleavage of all the different HA4B-5B substrates shown in Fig. 3. In case of the 3<sub>135F</sub> proteinase we found that only P1-Cys was accepted

**TABLE 1**  
Cleavage Efficiencies of Altered Proteinases at NS4B/5A Sites Carrying Different P1 Residues

Proteinase	P1			
	Cys	Leu	Ile	Val
3 <sub>154A</sub>	100 <sup>a</sup>	61	25	0
3 <sub>154V</sub>	100	104	48	21
3 <sub>154T</sub>	100	124	41	31
3 <sub>154F</sub>	100	0	0	0

<sup>a</sup> Relative amount of HA4B produced in the intracellular *trans*-cleavage assay during a 3-hr labeling period (given as percentages); values were determined by densitometry scanning and they represent mean values from three independent experiments (±5%). Cleavage at the P1-Cys residue was set at 100%.

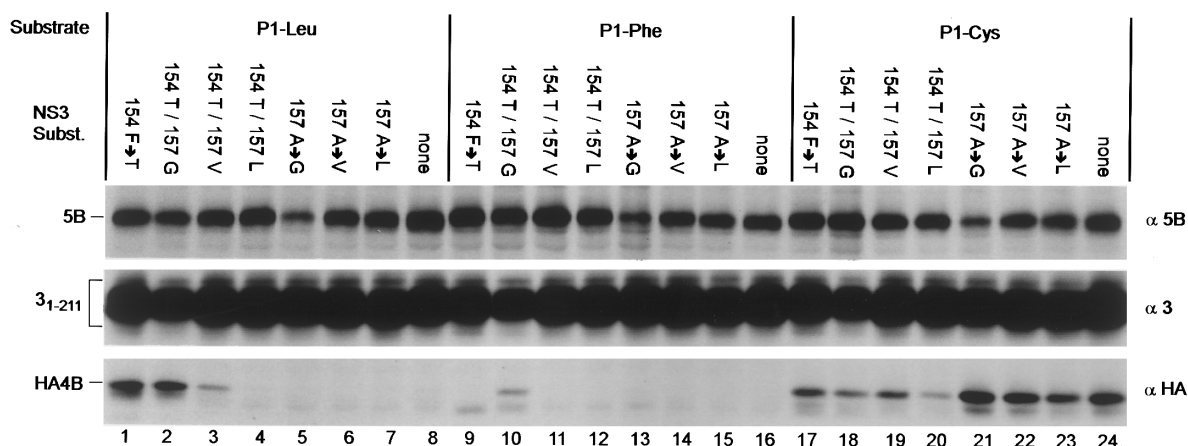


FIG. 5. Influence of Ala-157 on substrate specificity. Proteinases containing either single substitutions of residue 157 or double substitutions of residues 157 and 154 were coexpressed with NS4A<sub>2-54</sub> and HA4B-5B substrates carrying given P1 residues at the NS4B/5A site. As control, cleavage of the substrates by the parental NS3 proteinase was examined in parallel (lanes 8, 16, and 24). HCV-specific proteins were isolated by immunoprecipitation using antisera specified to the right and analyzed by SDS-PAGE. Stained gels were dried and exposed to X-ray films using an intensifying screen.

whereas the 3<sub>135G</sub> enzyme was completely inactive (not shown), suggesting that Leu-135 plays only a minor role in specificity and that a large aliphatic residue at this position is important for proteinase structure.

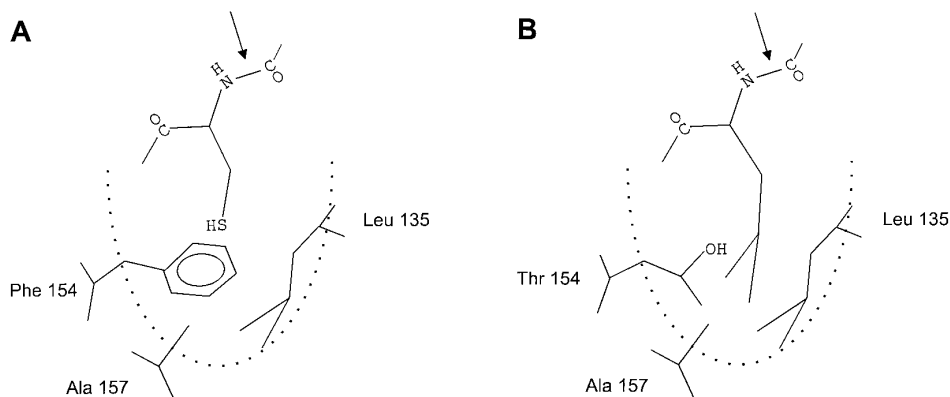
Thus far we had analyzed the role of Phe-154 and Leu-135 for substrate specificity. As described above the S1 pocket is formed by these two residues and Ala-157. In the experiment shown in Fig. 2 we found that substitutions of this alanine affected only cleavage efficiency, not specificity. However, since Ala-157 is "covered" by Phe-154, which forms the bottom of the S1 pocket, it was likely that substitutions at both positions would influence specificity synergistically and open up the SBP further. Therefore, double mutants were generated (3<sub>154T/157G</sub>, 3<sub>154T/157V</sub>, and 3<sub>154T/157L</sub>) and tested for cleavage of substrates with P1 residues with bulky nonpolar side chains in parallel with the parental NS3 proteinase and the corresponding single substitutions. As shown in Fig. 5, with respect to cleavage activity of the P1-Cys substrate, the enzymes carrying single substitutions of Ala-157 showed the following order: 3<sub>157G</sub> > 3<sub>157A</sub> (the parental proteinase) > 3<sub>157V</sub> > 3<sub>157L</sub> (lanes 21, 24, 22, and 23, respectively). This inverse correlation between size of the substituting residue and cleavage activity indicated a gradual alteration of the SBP affecting enzyme-substrate interaction. The analogous correlation was observed when the same Ala-157 substitutions were combined with the threonine substitution of Phe-154 and the double mutants were tested for cleavage of the P1-Cys substrate (lanes 17–20) or the P1-Leu substrate (lanes 1–4; the small amounts of HA-4B obtained with the 3<sub>154T/157G</sub> proteinase were due to low expression levels of this enzyme). In addition, the 3<sub>154T/157G</sub> enzyme had the broadest substrate specificity and could cleave even a P1-Phe substrate (lane 10) which was not accepted by the parental proteinase (lane 16) or any of the other mutants (lanes 9, 11–15). Thus, the simultaneous substitution of Phe-154 and Ala-157 by

threonine and glycine, respectively, had further altered specificity toward a proteinase accepting a very bulky hydrophobic P1 residue, indicating a further opening of the S1 pocket. A similar alteration of substrate specificity was described recently (Failla *et al.*, 1996).

## DISCUSSION

The aim of our study was a better understanding of the structural determinants of substrate specificity of the NS3 proteinase. Our finding that Phe-154 is the main determinant for substrate specificity of the NS3 proteinase agrees well with the model of the S1 pocket first described by Pizzi and co-workers in 1994. The X-ray crystal structure of NS3 shows that Phe-154 is located at the bottom of the S1 pocket, suggesting that this residue has a dual function: delimiting the length of the P1 residue that can fit the pocket and establishing a favorable interaction between the correct P1 residue cysteine and the aromatic side chain of the phenylalanine (Fig. 6A; Burley and Petsko, 1986). This dual function might be the explanation for the finding that cysteine is the most favorable P1 residue (and the only one found at the *trans*-cleavage sites). The only other residue for which we found cleavage was alanine, but processing efficiency was very low. It is possible that alanine fits the SBP by size but barely interacts with the aromatic ring of Phe-154. In a recent report Failla and co-workers (1996) described a double mutant in which Phe-154 and Ala-157 were replaced by threonine and glycine, respectively. When tested for cleavage of an NS5A/5B substrate, they found that the enzyme accepted, in addition to the P1-Cys, a P1-Phe residue. While this finding agrees with our results described here for the NS4B/5A site, we performed a more extensive analysis and showed that Ala-157 plays a minor role in specificity. As we know now from the crystal structure, Phe-154 forms the bottom of





**FIG. 6.** Hypothetical model of the SBP of the parental NS3 proteinase (A) and the proteinase carrying a single threonine substitution at position 154 (B). The S1 pocket is indicated with the dotted line and it harbors the P1 side chain of cysteine (A) or leucine (B). The arrow indicates the bond cleaved by the proteinase. The model shown in (A) is based on suggestions from Pizzi *et al.* (1994) and Failla *et al.* (1996) and it includes an interpretation of the substrate modeling described by Love *et al.* (1996).

the SBP covering Ala-157 located underneath (Fig. 6A). Consequently, in agreement with our mutation analysis alterations of Ala-157 affect specificity only when Phe-154 is also changed (Fig. 6B). This structure also explains the inverse correlation between an increase in size of the amino acid at position 157 and the reduction of efficiency with which substrates with P1 residues of increasing sizes are cleaved. While this model explains the roles of Phe-154 and Ala-157 for specificity, the function of Leu-135 is less clear. A drastic change in the size of this residue (135 L → G) resulted in a completely inactive proteinase whereas a moderate change (135 L → F) did not impair cleavage activity or specificity, suggesting that Leu-135 may play a structural role.

For all serine proteinases, amino acids forming the catalytic triad are found in a particular geometric arrangement in which the serine side chain is hydrogen-bonded to the histidine imidazole, which in turn is hydrogen-bonded to the carboxylate of aspartate (for a review see Kraut, 1977). This conformation ensures that the negative charge on the carboxylate makes the serine oxygen nucleophilic and in this way reactive toward the substrate. For HCV NS3 the catalytic triad is formed by His-57, Asp-81, and Ser-139. However, in the crystal structure described by Love and co-workers (1996), Asp-81 appears rotated away from His-57 and stabilized by an ion pair to Arg-155. We found that substitution of Arg-155 by serine had no effect on cleavage efficiency, indicating that the ion pair is not essential for activity or specificity. In fact, as discussed by the authors minor modifications of structural orientations of amino acids at positions 80–82 bring Asp-81 closer to His-57. This would create a classic catalytic triad and it was suggested that the positional deviation of Asp-81 observed in this structure might be due to the lack of NS4A or the crystallization process.

We found that all substitutions at position 154 created enzymes preferring substrates with P1-Cys or P1-Leu residues. The only amino acids with a branched side chain for which cleavage was detected were isoleucine

and valine, indicating that a methyl group attached to the C $\beta$  of the amino acid side chain constitutes a steric hindrance impairing cleavage activity. Although the natural P1 residue is a cysteine—at least at the *trans* cleavage sites—a threonine is also accepted in the context of the NS3/4A (*cis*) cleavage site.

Several studies have identified amino acids determining substrate specificities of cellular serine proteinases. In the case of trypsin, Ala-157 and Val-167 of HCV NS3 correspond to two glycine residues at positions 216 and 226 (chymotrypsin numbering; Fig. 1A). The trypsin–trypsin inhibitor structure suggests that these glycine residues shape the pocket to accommodate P1 residues with long side chains. Asp-189 at the bottom of the SBP favors electrostatic interactions with positively charged amino acids, explaining the specificity of trypsin for basic P1 residues. The S1-binding sites of trypsin and chymotrypsin are nearly identical in structure and primary sequence (Fig. 1A). However, the aspartic acid at the bottom of the pocket is missing, creating a rather large hydrophobic pocket accommodating P1 residues with large hydrophobic side chains. In this respect, the specificity of the NS3<sub>154T/157G</sub> proteinase was altered toward the specificity of chymotrypsin. In the case of elastase, valine and threonine are found at positions 216 and 226 (Fig. 1A). These bulkier residues narrow the SBP, generating a preference for small nonpolar P1 side chains. From this point of view, elastase most closely resembles the NS3 proteinase at these positions.

Numerous studies analyzed amino acid exchanges in the context of the different cleavage sites of the polyprotein (Kolykhalov *et al.*, 1994; Komoda *et al.*, 1994; Bartenschlager *et al.*, 1995a) and found differences in the sensitivities toward P1 substitutions, suggesting that the S1 pocket is not the sole determinant of substrate specificity. Substrate modeling studies indicate that interaction of the NS3 proteinase with its substrate is accomplished by several synergistic factors: (i) specific binding of the P1 residue to the S1 pocket; (ii) stabilization of this bind-

ing by a favorable interaction between the aromatic ring of Phe-154 and the sulfhydryl group of the P1-Cys or, to a lesser extent, the hydroxyl group of threonine; (iii) a continuous P2-P6 main chain interaction; and (iv) a possible interaction between the acidic P6 residue of the substrate and Arg-161 or Lys-165 (Pizzi *et al.*, 1994; Kim *et al.*, 1996; Love *et al.*, 1996). In agreement with these results Urbani and co-workers (1997) provided evidence that ground state binding of the substrate is achieved by multiple interactions involving distal residues whereas the P1 residue is the primary determinant of efficiency with which the bound substrate will proceed through the transition state. In this respect the S1 pocket appears to be a primary determinant for catalysis.

In summary we have analyzed the different contributions the residues forming the SBP have for specificity and activity. In agreement with other reports describing the alteration of substrate specificity of proteinases by distinct amino acid substitutions (e.g., Bone *et al.*, 1989; Hedstrom *et al.*, 1992; Failla *et al.*, 1996), we found that specificity can be restricted to a few determinants within the enzyme. However, given the elaborate interactions between the NS3 proteinase and the residues around the scissile bonds, further studies will be required to elucidate how fine specificity of the enzyme is brought about and how to design efficient inhibitors of the NS3 proteinase of HCV.

## ACKNOWLEDGMENTS

We thank Toni Krohn for initial advice on identifying amino acids important in substrate specificity and Volker Lohmann for critical discussions. We are grateful for excellent technical assistance provided by Ulrike Herian. We also thank B. Moss for the gift of plasmid pTM1 and the recombinant vaccinia virus vTF7-3. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (BA 1505/1-2).

## REFERENCES

- Bartenschlager, R. (1997). Molecular targets in inhibition of hepatitis C virus replication. *Antiviral Chem. Chemother.* **8**, 281–301.
- Bartenschlager, R., Ahlborn-Laake, L., Yasargil, K., Mous, J., and Jacobsen, H. (1995a). Substrate determinants for cleavage in *cis* and *trans* by the hepatitis C virus NS3 proteinase. *J. Virol.* **69**, 198–205.
- Bartenschlager, R., Ahlborn-Laake, L., Mous, J., and Jacobsen, H. (1993). Nonstructural protein 3 of the hepatitis C virus encodes a serin-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *J. Virol.* **67**, 3835–3844.
- Bartenschlager, R., Ahlborn-Laake, L., Mous, J., and Jacobsen, H. (1994). Kinetic and structural analysis of hepatitis C virus polyprotein processing. *J. Virol.* **68**, 5045–5055.
- Bartenschlager, R., Lohmann, V., Wilkinson, T., and Koch, J. O. (1995b). Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation. *J. Virol.* **69**, 7519–7528.
- Bazan, J. F., and Fletterick, R. J. (1988). Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: Structural and functional implications. *Proc. Natl. Acad. Sci. USA* **85**, 7872–7876.
- Behrens, S.-E., Tomei, L., and De Francesco, R. (1996). Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.* **15**, 12–22.
- Bone, R., Silen, J. L., and Agard, D. A. (1989). Structural plasticity broadens the specificity of an engineered protease. *Nature* **339**, 191–195.
- Bordo, D., and Argos, P. (1991). Suggestions for “safe” residue substitutions in site-directed mutagenesis. *J. Mol. Biol.* **217**, 721–729.
- Burley, S. K., and Petsko, G. A. (1986). Amino-aromatic interactions in proteins. *FEBS Lett.* **203**, 139–143.
- Cuthbert, J. A. (1994). Hepatitis C: Progress and problems. *Clin. Microbiol. Rev.* **7**, 505–532.
- Dubuisson, J., Hsu, H. H., Cheung, R. C., Greenberg, H. B., Russell, D. G., and Rice, C. M. (1994). Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and sindbis viruses. *J. Virol.* **68**, 6147–6160.
- Eckart, M. R., Selby, M., Masiarz, F., Lee, C., Berger, K., Crawford, C., Kuo, C., Kuo, G., Houghton, M., and Choo, Q.-L. (1993). The hepatitis C virus encodes a serin proteinase involved in processing of the putative nonstructural proteins from the viral polyprotein precursor. *Biochem. Biophys. Res. Commun.* **192**, 399–406.
- Failla, C., Tomei, L., and De Francesco, R. (1995). An amino-terminal domain of the hepatitis C virus NS3 protease is essential for interaction with NS4A. *J. Virol.* **69**, 1769–1777.
- Failla, C. M., Pizzi, E., De Francesco, R., and Tramontano, A. (1996). Redesigning the substrate specificity of hepatitis C virus NS3 proteinase. *Folding Design* **1**, 35–42.
- Failla, C., Tomei, L., and De Francesco, R. (1994). Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *J. Virol.* **68**, 3753–3760.
- Fuerst, T. R., Niles, E. G., Studier, F. W., and Moss, B. (1986). Eucaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA* **83**, 8122–8126.
- Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., and Rice, C. M. (1993b). Characterization of the hepatitis C virus-encoded serine proteinase: Determination of proteinase-dependent polyprotein cleavage sites. *J. Virol.* **67**, 2832–2843.
- Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., and Rice, C. M. (1993a). A second hepatitis C virus-encoded proteinase. *Proc. Natl. Acad. Sci. USA* **90**, 10583–10587.
- Hedstrom, L., Szilagyi, L., and Rutter, W. J. (1992). Converting trypsin to chymotrypsin: the role of surface loops. *Science* **255**, 1249–1253.
- Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K., and Shimotohno, K. (1993). Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus. *J. Virol.* **67**, 4665–4675.
- Houghton, M. (1996). Hepatitis C viruses. In “Virology” (B. N. Fields, P. M. Knipe, P. M. Howley, *et al.*, Eds.), pp. 1035–1058. Lippincott–Raven, Philadelphia.
- Hsu, H. H., Donets, M., Greenberg, H. B., and Feinstone, S. M. (1993). Characterization of hepatitis C virus structural proteins with a recombinant baculovirus expression system. *Hepatology* **17**, 763–771.
- Inchauspe, G., Yebedee, S., Lee, D.-H., Sugitani, M., Nasoff, M., and Prince, A. (1991). Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates. *Proc. Natl. Acad. Sci. USA* **88**, 10292–10296.
- Kim, D. W., Gwack, Y., Han, J. H., and Choe, J. (1995). C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase. *Biochem. Biophys. Res. Commun.* **215**, 160–166.
- Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., Chambers, S. P., Markland, W., Lepre, C. A., O'Malley, E. T., Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R., and Thomson, J. A. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* **87**, 343–355.
- Koch, J. O., Lohmann, V., Herian, U., and Bartenschlager, R. (1996). *In vitro* studies on the activation of the hepatitis C virus NS3 proteinase by the NS4A cofactor. *Virology* **221**, 54–66.
- Kolykhalov, A. A., Agapov, E. V., and Rice, C. M. (1994). Specificity of the hepatitis C virus NS3 protease: Effects of substitutions at the

- 3/4A, 4A/4B, 4B/5A, and 5A/5B cleavage sites on polyprotein processing. *J. Virol.* **68**, 7525–7533.
- Komoda, Y., Hijikata, M., Sato, S., Asabe, S.-I., Kimura, K., and Shimotohno, K. (1994). Substrate requirements of hepatitis C virus serine proteinase for intermolecular polypeptide cleavage in *Escherichia coli*. *J. Virol.* **68**, 7351–7357.
- Kraut, J. (1977). Serine proteases: structure and mechanism of catalysis. *Annu. Rev. Biochem.* **46**, 331–358.
- Leinbach, S. S., Bhat, R. A., Xia, S.-M., Hum, W.-T., Stauffer, B., Davis, A. R., Hung, P. P., and Mizutani, S. (1994). Substrate specificity of the NS3 serine proteinase of hepatitis C virus as determined by mutagenesis at the NS3/NS4A junction. *Virology* **204**, 163–169.
- Lin, C., Pragai, B. M., Grakoui, A., Xu, J., and Rice, C. M. (1994). Hepatitis C virus NS3 serine proteinase: *trans*-cleavage requirements and processing kinetics. *J. Virol.* **68**, 8147–8157.
- Lin, C., Thomson, J. A., and Rice, C. M. (1995). A central region in the hepatitis C virus NS4A protein allows formation of an active NS3-NS4A serine proteinase complex *in vivo* and *in vitro*. *J. Virol.* **69**, 4373–4380.
- Lohmann, V., Körner, F., Herian, U., and Bartenschlager, R. (1997). Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J. Virol.*, in press.
- Love, R. A., Parge, H. E., Wickersham, J. A., Hostomsky, Z., Habuka, N., Moomaw, E. W., Adachi, T., and Hostomska, Z. (1996). The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell* **87**, 331–342.
- Matsumoto, M., Hwang, S. B., Jeng, K.-S., Zhu, N., and Lai, M. M. (1996). Homotypic interaction and multimerization of hepatitis C virus core protein. *Virology* **218**, 43–51.
- Matsuura, Y., Harada, S., Suzuki, R., Watanabe, Y., Inoue, Y., Saito, I., and Miyamura, T. (1992). Expression of processed envelope protein of hepatitis C virus in mammalian and insect cell. *J. Virol.* **66**, 1425–1431.
- Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D. (Eds.). (1995). Classification and Nomenclature of Viruses: Sixth Report of the International Committee on Taxonomy of Viruses pp. 424–426. Springer Verlag, Vienna/New York.
- Pizzi, E., Tramontano, A., Tomei, L., LaMonica, N., Failla, C., Sardana, M., Wood, T., and DeFrancesco, R. (1994). Molecular model of the specificity pocket of the hepatitis C virus protease: Implications for substrate recognition. *Proc. Natl. Acad. Sci. USA* **91**, 888–892.
- Ralston, R., Thudium, K., Berger, K., Kuo, C., Gervase, B., Hall, J., Selby, M., Kuo, G., Houghton, M., and Choo, Q.-L. (1993). Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. *J. Virol.* **67**, 6753–6761.
- Rice, C. M. (1996). Flaviviridae: The viruses and their replication. In "Virology" (B. N. Fields, P. M. Knipe, P. M. Howley, *et al.*, Eds.), pp. 931–959. Lippincott-Raven, Philadelphia.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Santolini, E., Migliaccio, G., and La Monica, N. (1994). Biosynthesis and biochemical properties of the hepatitis C virus core protein. *J. Virol.* **68**, 3631–3641.
- Satoh, S., Tanji, Y., Hijikata, M., Kimura, K., and Shimotohno, K. (1995). The N-terminal region of hepatitis C virus nonstructural protein 3 (NS3) is essential for stable complex formation with NS4A. *J. Virol.* **69**, 4255–4260.
- Schechter, J., and Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162.
- Spaete, R. R., Alexander, D., Rugroden, M., Choo, Q.-L., Berger, K., Crawford, K., Kuo, C., Leng, S., Lee, C., Ralston, R., Thudium, K., Tung, J. W., Kuo, G., and Houghton, M. (1992). Characterization of the hepatitis C virus E2/NS1 gene product expressed in mammalian cells. *Virology* **188**, 819–830.
- Suzich, J. A., Tamura, J. K., Palmer-Hill, F., Warren, P., Grakoui, A., Rice, C. M., Feinstone, S. M., and Collett, M. S. (1993). Hepatitis C virus NS3 protein polynucleotide-stimulated nucleoside triphosphatase and comparison with related pestivirus and flavivirus enzymes. *J. Virol.* **67**, 6152–6158.
- Takamizawa, A., Mori, C., Fuke, I., Manabe, S., Murakami, S., Fujita, J., Onishi, E., Andoh, T., Yoshida, I., and Okayama, H. (1991). Structure and organization of the hepatitis C virus genome isolated from human carriers. *J. Virol.* **65**, 1105–1113.
- Tanji, Y., Hijikata, M., Satoh, S., Kaneko, T., and Shimotohno, K. (1995). Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. *J. Virol.* **69**, 1575–1581.
- Tomei, L., Failla, C., Santolini, E., De Francesco, R., and La Monica, N. (1993). NS3 is a serine protease required for processing of hepatitis C virus polyprotein. *J. Virol.* **67**, 4017–4026.
- Urbani, A., Bianchi, E., Narjes, F., Tramontano, A., De Francesco, R., Steinkühler, C., and Pessi, A. (1997). Substrate specificity of the hepatitis C virus serine protease NS3. *J. Biol. Chem.* **272**, 9204–9209.